

⑫ 公開特許公報(A)

昭64-3700

⑤ Int. Cl.⁴G 10 L 3/00
G 06 F 3/16

識別記号

3 0 1
3 2 0

庁内整理番号

Z-8842-5D
H-7341-5B

⑬ 公開 昭和64年(1989)1月9日

審査請求 未請求 発明の数 3 (全7頁)

⑭ 発明の名称 音声認識方法および装置

⑮ 特 願 昭62-157625

⑯ 出 願 昭62(1987)6月26日

⑰ 発 明 者 麻 生 隆 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

⑱ 出 願 人 キヤノン株式会社 東京都大田区下丸子3丁目30番2号

⑲ 代 理 人 弁理士 谷 義 一

明 細 書

1. 発明の名称

音声認識方法および装置

2. 特許請求の範囲

1) 第1の音声情報を入力し、

前記第1の音声情報の認識に用いられる第2の音声情報を記録媒体に予め記録し、

その記録された第2の音声情報を読み取り、

その読み取られた第2の音声情報および第1の音声情報を送信し、

その送信されてきた第1の音声情報および第2の音声情報を受信し、

その受信された第2の音声情報を記憶し、

当該記憶された第2の音声情報に基づいて前記受信された第1の音声情報の音声認識を行う

ことを特徴とする音声認識方法。

2) 第1の音声情報を入力する入力手段と、

前記第1の音声情報の認識に用いられる第2の音声情報を予め記録した記録媒体と、

該記録媒体に記録された前記第2の音声情報を読み取る読み取り手段と、

該読み取り手段により読み取られた第2の音声情報および前記入力手段に入力された前記第1の音声情報を送信する送信手段と

を具えたことを特徴とする音声認識装置。

3) 第1の音声情報および該第1の音声情報の認識に用いられる第2の音声情報を受信する受信手段と、

該受信手段により受信された前記第2の音声情報を記憶する記憶手段と、

当該記憶された前記第2の音声情報に基づいて前記受信手段により受信された前記第1の音声情報の音声認識を行う認識手段と

を具えたことを特徴とする音声認識装置。

(以下、余白)

3. 発明の詳細な説明

〔産業上の利用分野〕

本発明は、音声を認識する音声認識方法および装置に関する。

〔従来の技術〕

一般に、電話から入力した音声を認識するシステムが知られている。これらシステムの多くは予め送話者の音声から抽出したもしくは音声の特徴パターンをパラメータとしてメモリに登録されており、音声認識回路がメモリに記憶された特徴パターンと電話から送られてくる送話者の音声（単語）を比較し、送話者が送る音声（単語）と同じ音声の特徴パターンを持つものを抽出することにより、音声の意味する単語を認識している。

〔発明が解決しようとする問題点〕

ところが、このようなシステムを多くの人が使用する場合、各話者ごとに音声を音声認識回路に登録する作業が必要であり、話者の数が増えると各話者の音声を記憶しておくメモリ容量が膨大になってしまうという問題点があった。

する入力手段と、第1の音声情報の認識に用いられる第2の音声情報を予め記録した記録媒体と、記録媒体に記録された第2の音声情報を読み取る読み取り手段と、読み取り手段により読み取られた第2の音声情報および入力手段に入力された第1の音声情報を送信する送信手段とを具えたことを特徴とする。

本発明の第3の形態は、第1の音声情報および第1の音声情報の認識に用いられる第2の音声情報を受信する受信手段と、受信手段により受信された第2の音声情報を記憶する記憶手段と、記憶された第2の音声情報に基づいて受信手段により受信された第1の音声情報の音声認識を行う認識手段とを具えたことを特徴とする。

〔作 用〕

本発明の第1の形態においては、第1の音声情報を入力し、第1の音声情報の認識に用いられる第2の音声情報を記録媒体に予め記録し、その記録された第2の音声情報を読み取り、その読み取られた第2の音声情報および第1の音声情報を送

不特定話者についての音声認識を行う回路も考えられているが、音声の認識確率が低く、未だ実用化に到っていないのが現状である。

そこで、本発明の目的は、このような問題点を解決し、大多数の不特定話者が発生する音声を確実に認識できる音声認識方法および装置を提供することにある。

〔問題点を解決するための手段〕

このような目的を達成するために、本発明の第1の形態は、第1の音声情報を入力し、第1の音声情報の認識に用いられる第2の音声情報を記録媒体に予め記録し、その記録された第2の音声情報を読み取り、その読み取られた第2の音声情報および第1の音声情報を送信し、その送信されてきた第1の音声情報および第2の音声情報を受信し、その受信された第2の音声情報を記憶し、記憶された第2の音声情報に基づいて受信された第1の音声情報の音声認識を行うことを特徴とする。

本発明の第2の形態は、第1の音声情報を入力

し、その送信されてきた第1の音声情報および第2の音声情報を受信し、その受信された第2の音声情報を記憶し、記憶された第2の音声情報に基づいて受信された第1の音声情報の音声認識を行うことを特徴としているので、音声認識に用いる第2の音声情報を予め記憶しておく必要がなく、不特定多数の音声入力者から入力される第1の音声情報を確実に認識することができる。

本発明の第2の形態においては、予め記録媒体に記録された音声認識に用いられる第2の音声情報を読み取り手段により読み取って第1の音声情報を送信するようにしたので、受信側ではこの第2の音声情報に基づいて第1の音声情報を認識することができる。

本発明の第3の形態においては、第1の音声の受信に先立ってこの第1の音声の音声認識に用いられる音声情報をも受信して記憶するようにしたので、この第2の音声情報に基づいて第1の音声の情報を受信したときに第2の音声を認識することができ、受信側において受信する音声情報のそれ

それに対応する音声認識するための第1の音声情報を前もって記憶しておく必要がない。

[実施例]

以下に、図面を参照して本発明の一実施例を詳細に説明する。

第1図は本発明実施例における構成の一例を示す。

第1図において、11はカードリーダーであり、このカードリーダー11は磁気カードや光のカードを読み取る一般的に知られているものを使用することができる。カードに予め書き込まれた第2の音声情報としての音声における特徴パターンを読み取る。なお、上記カードには、音声の特徴パターンとして使用するための予め定められた文字列に対応する音声情報が記録されている。12は制御部であり、制御部12はカードリーダー11により読み取られた音声の特徴パターンを公衆回線に対して出力するように制御する。

13は第1の音声情報としての音声の入出力を行う受話器であり、14はカードリーダー11から読み取

の一例を示す。

第2図において、15はカードリーダー11により読み取られた音声の特徴パターンに関する信号を演算処理装置(CPU)16に転送するインタフェースである。16は本発明に係わる第4図示の制御手順を実行するCPUである。17は第4図示の制御手順の他、公衆回線へ情報を転送するための制御手順を記憶するメモリである。18はCPU16に送受信される情報を転送するインタフェースである。

19はモデムと呼ばれるアナログデジタル(A/D)変換・デジタルアナログ(D/A)変換器である。CPU16が処理する信号はデジタル信号であり、公衆回線が伝える信号はアナログ信号であるので、モデム19により信号の入出力に応じて信号をA/D変換もしくはD/A変換を行う。

第3図は本発明実施例におけるカードデータ格納部21と音声認識処理部の構成の一例を示す。

第3図において、24はモデムであり、25はインタフェースである。26はCPUであり、CPU26は第5図の制御手順(ステップS24→S28)の処理を実

行われた音声の特徴パターンおよび受話器と送受信する音声情報を択一的に公衆回線に接続する切り換え部である。

20は公衆回線から送られてくる情報をその種類に応じてカードデータ格納部21もしくは音声認識部22へ切り換え接続する切り換え部である。21は音声の特徴パターンを記憶しておくカードデータ格納部である。

22は受話器13から送られていく音声情報の意味を音声の特徴パターンに基いて認識する音声認識部である。なお、この音声の特徴パターンはカードデータ格納部21に記憶された音声の特徴パターンが用いられる。

23は情報処理機器などの電子機器であり、情報処理機器は音声認識部22により認識されたメッセージ内容に対応した処理を行う。

動作説明に先立って、このような構成における制御部12、カードデータ格納部21および音声認識部22の具体的な構成例について説明する。

第2図は本発明実施例における制御部12の構成

を行う。27は公衆回線を介して送られてきた音声の特徴パターンを記憶しておく書き換え自在のメモリである。

28は受話器13から送られてきた音声信号をA/D変換するA/D変換器である。29はインタフェースである。30はデジタル変換された音声信号を記憶しておくメモリである。

31はCPUであり、CPU31はカードデータ格納部21のメモリ27に格納された音声の特徴パターンに基いて、メモリ30に格納された音声の認識処理(第5図ステップS30→S34)を行う。

32はCPU31から送られる制御信号を情報機器23に転送するインタフェースである。

次に、このような構成における本実施例の動作を第4図および第5図のフローチャートを参照しながら説明する。

第4図は送信側の処理手順の一例を示す。

第4図において、装置利用者がカードをカードリーダー11に挿入する。そして次に、利用者が受信側の電話番号を受話器13によりダイヤルし、送信

側と受信側の接続を行う（ステップS13～S14）。制御部12が受信側との公衆回線の接続を確認すると、制御部12から切り換え指示信号を切り換え部14に送り、公衆回線と制御部12とを接続する（ステップS15）。

次に、制御部12はカードリーダー11にカードから音声の特徴パターンを読み取るように指示する。すると、カードリーダー11は、カードから読み取った音声の特徴パターンを制御部12へ送信する（ステップS16）。

制御部12においては、この音声の特徴パターンをCPU16が受信すると、CPU16は音声の特徴パターンをメモリ17に格納する（ステップS17）。次に、CPU16はメモリ17に記憶されている音声特徴パターンに関するデータを、公衆回線を転送するための通信コード体系に従って符号化を行い、モデム18より公衆回線に出力する（ステップS19）。

なお、CPU16は受信側から送られてくる送信情報の受取りを確認する制御信号を受信し、音声の特徴パターンの通信にエラーが生じたことを検出

をカードデータ格納部21のCPU26がメモリ27に記憶する（ステップS27）。

なお、CPU26は入力情報のコード形態をバリディチェックなどにより確認し、入力情報に異常が生じたときはエラー情報を送信側に送り、ステップS26から情報の受信をやり直す。

入力情報が正常なときは、CPU26は音声認識部22のCPU31に制御を渡す（ステップS28）。

CPU31は、公衆回線を音声認識部22と接続するように切り換え部20に指示し、送信側から送られてくる送話者のメッセージを受信する（ステップS29～S30）。

なお、このメッセージはCPU31によりメモリ30に記憶される。上記メッセージの受信が終了するまで、CPU31は、メモリ27に格納された音声の特徴パターンに基づいて、送話者のメッセージが何であるかを認識し、認識結果に対応した情報処理を、例えば、機器の作動や中止を情報処理機器23へ指示する（ステップS31～S33）。送信側との通話が終了すると、CPU31は公衆回線を遮断し、本

したときは、ステップS18に戻り、メモリ17に記憶してある音声の特徴パターンの送信をやり直す。

音声の特徴パターンの送信が正常に終了したことをCPU16が確認すると（ステップS20）、CPU16は切り換え部14に対して、公衆回線の接続を受話器13へ切り換えるように指示し、切り換え部14により公衆回線の接続を切り換える（ステップS21）。

以後、操作者が受話器13により音声により必要なメッセージを受信側に送信すると、CPU16は、公衆回線を遮断し、本制御手順を終了する（ステップS22～S23）。

次に、受信側の情報処理について第5図を用いて説明する。

第5図において、受信側において公衆回線との接続がなされると（ステップS24）、音声認識部22は、切り換え部20に指示を行い、公衆回線とカードデータ格納部21を接続する（ステップS25）。

次に、送信側から送られてくる音声の特徴パタ

制御手順を終了する。

このように、受信側では送信側から最初に送られてくる音声の特徴パターンを書き換え自在なメモリ27に記憶しておき、その後送話者が送るメッセージの内容をメモリ27に記憶してある音声の特徴パターンに基づき認識するので、不特定多数の送信側と交信する場合でも送話者の音声の特徴パターンのそれぞれを予め受信側のメモリに登録しておく必要がなくなる。

なお、本実施例の利用形態として、音声の特徴パターンはカードを使用する利用者が例えば予めマイクなどから“1”～“10”までの音声を入力し、入力された音声データを磁気記録カードや光記録カードに書き込んでおく。そして本実施例において、送話者が“1”～“10”などのメッセージを受信側に送ると、受信側では最初に送信されてきた“1”～“10”の音声の特徴パターンに基づいて、送話者が発生する音声が“1”～“10”のいずれの数字かを識別することになる。一例として音声の特徴パターンとして数字を説明したが、

数字に限らず、合い言葉を用いても可能である。

従って、受信側では、カードを使用している送話者の音声のカードに登録された音声と一致しているか否かをとも確認することが可能であり、カードの不正利用を防ぐことができる。この不正利用の防止率は、キーワードなどをキーボードから入力してカードの持ち主とシステム利用者が一致していることを確認するシステムよりもずっと高くなる。

次に、音声の認識方法については音声の周波数分析を行い、音声の特徴パターンについての周波数分析結果との比較を行うDPマッチング（動的計画法）などが知られており、さらに、この比較を行う処理をLSI（集積回路）化した演算処理回路が知られているので、このLSIを用いると受信側の装置を小型化できる。

さらに、音声の特徴パターンに使用する特徴パラメータとして、音声のピッチ、一定期間ごとの周波数スペクトルの時間的な推移、ホルマントの

位置の推移などが考えられるが、これは音声認識部22で認識する対象が何であるか（例えば限定された単語単位での認識、単語単位での認識など）により、その認識に最も必要なパラメータの抽出を行うべきである。

〔発明の効果〕

以上説明したように、本発明によれば、カードに音声の特徴パターンを登録しておき、電話での音声認識の前にカードに登録した音声の特徴パターンを音声認識装置に送り、その音声の特徴パターンを用いて送話者の音声認識を行うことにより、不特定多数の送話者の音声をより確実に認識でき、さらには、音声の特徴パターンを受信側で予め記憶しておく必要もないので、音声認識のための音声情報を記憶しておくメモリのメモリ容量を小さくでき、以って、装置を小型化できるという効果が得られる。

4. 図面の簡単な説明

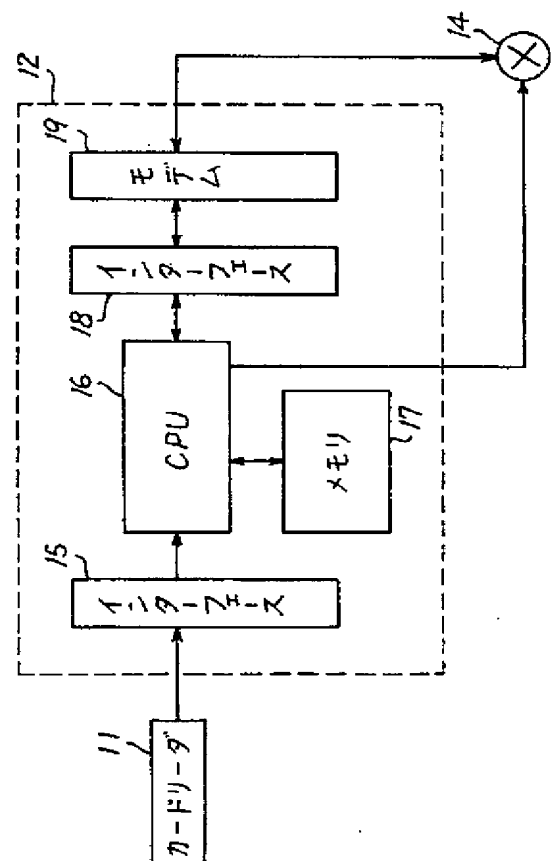
第1図は本発明実施例におけるシステム構成の一例を示すブロック図、

第2図は本発明実施例の制御部12の構成の一例を示すブロック図、

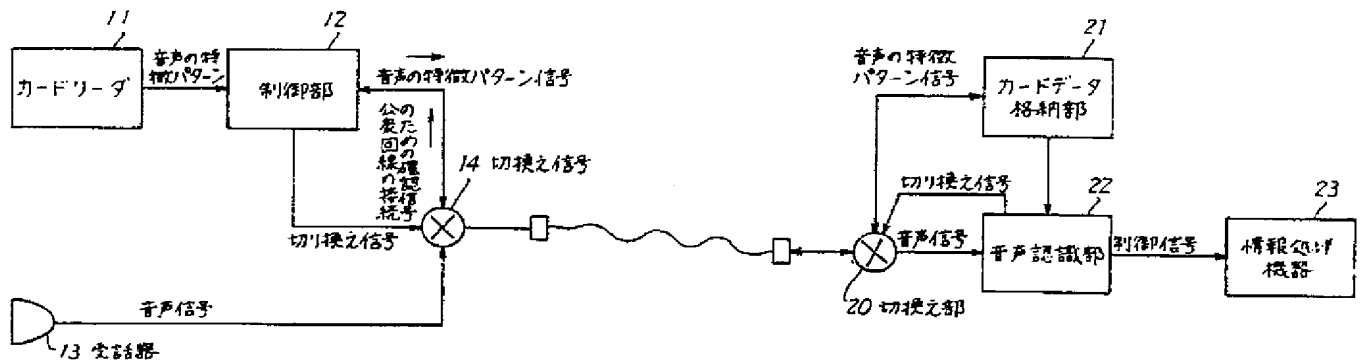
第3図は本発明実施例のカードデータ格納部21および音声認識部22の構成の一例を示すブロック図、

第4図および第5図は本発明実施例における動作手順の一例を示すフローチャートである。

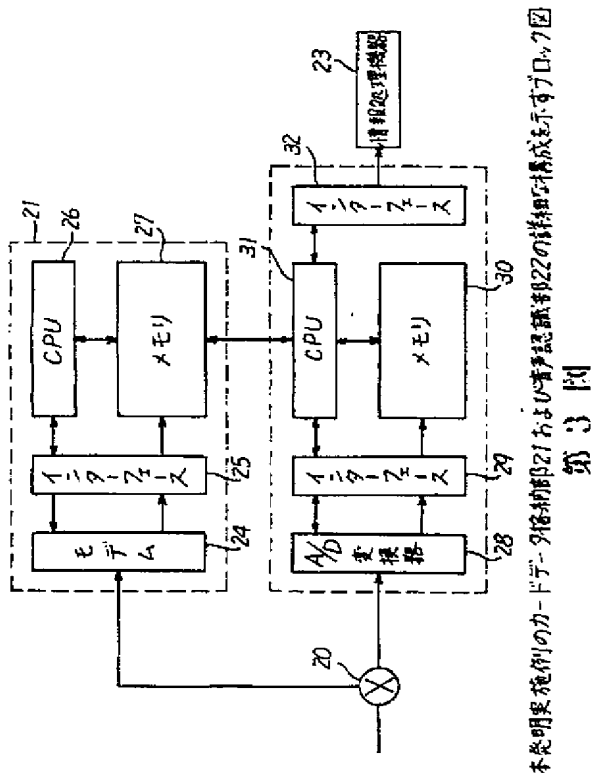
- 11…カードリーダー、
- 12…制御部、
- 13…受話器、
- 14, 20…切り換え部、
- 21…カードデータ格納部、
- 22…音声認識部、
- 23…情報処理機器。



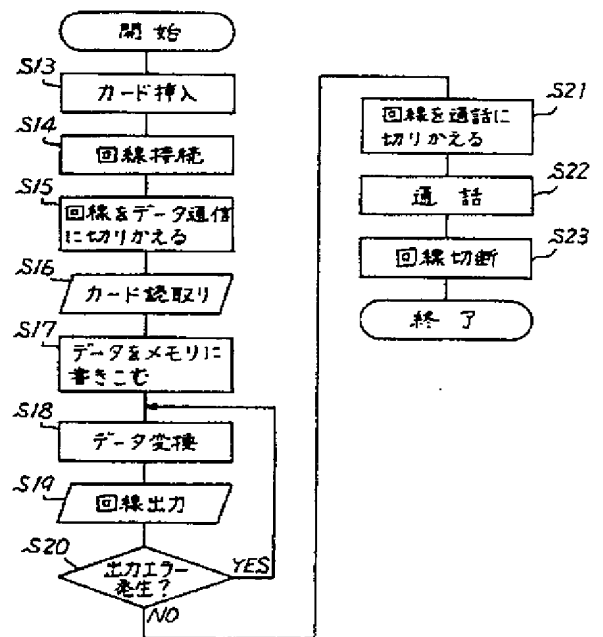
本発明実施例の制御部12の構成を示すブロック図
第2図



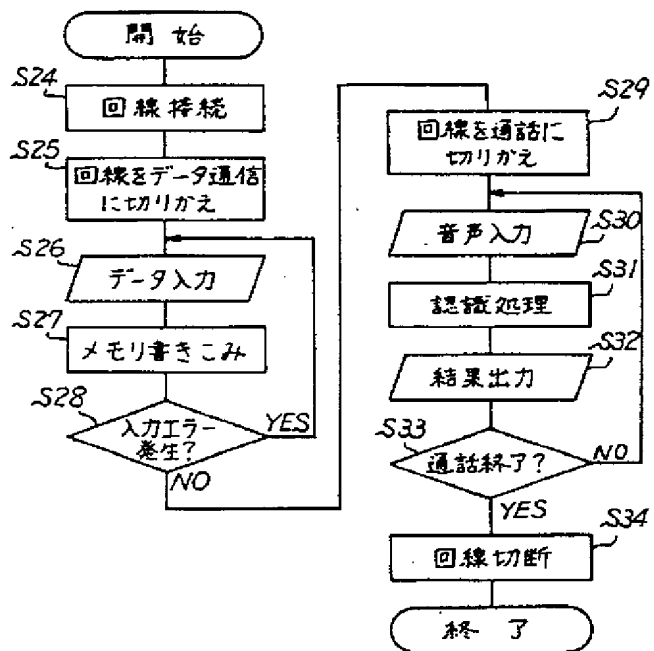
本発明実施例のブロック図
第 1 図



本発明実施例のカードデータ格納部21および音声認識部22の詳細構成を示すブロック図
第 3 図



本発明実施例のフローチャート
第 4 図



本発明実施例のフローチャート
第 5 図